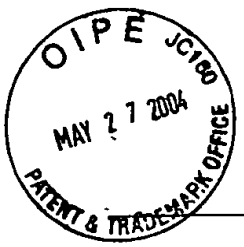
AF/2655/$

PATENT



Case Docket No. AMAZON.059A
Date: May 24, 2004
Page 1

D/EXT
Brief
6/4/04
OK

In re application of	:	Nicholas J. Lee
Appl. No.	:	09/650,173
Filed	:	August 29, 2000
For	:	VOICE INTERFACE AND METHODS FOR IMPROVING RECOGNITION ACCURACY OF VOICE SEARCH QUERIES
Examiner	:	Michael N. Opsasnick
Art Unit	:	2655

I hereby certify that this correspondence and all marked attachments are being deposited with the United States Postal Service as first-class mail in an envelope addressed to: Mail Stop Appeal Brief - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on

May 24, 2004
(Date)

Ronald J. Schoenbaum, Reg. No. 38,297

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Transmitted herewith in triplicate is an Appellant's Brief to the Board of Patent Appeals:

Enclosed are the following:

(X) A check in the amount of $440.00 to cover the filing fees; and
(X) A return prepaid postcard.

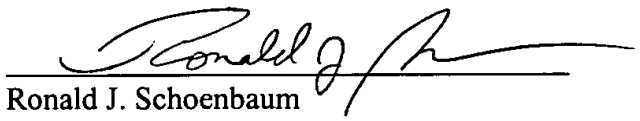
An extension of time to respond for 1 month is hereby requested.

Time Extension Fee:

(X) one month ($110)
() two months ($420)
() three months ($950)

05/28/2004 EFLORES 00000176 09650173
02 FC:1251 110.00 OP

If applicant has not requested a sufficient extension of time and/or has not paid any other fee in a sufficient amount to prevent the abandonment of this application, please consider this as a Request for an Extension for the required time period and/or authorization to charge our Deposit Account No. 11-1410 for any fee which may be due. Please credit any overpayment to Deposit Account No. 11-1410.



Ronald J. Schoenbaum
Registration No. 38,297
Attorney of Record
Customer No. 20,995
(949) 760-0404

H:\DOCS\ROS\ROS-6261.DOC\dns
042304

AMAZON.059A

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	:	Nicholas Lee	Group Art Unit: 2655
Appl. No.	:	09/650,173	
Filed	:	August 29, 2000	
For	:	VOICE INTERFACE AND METHODS FOR IMPROVING RECOGNITION ACCURACY OF VOICE SEARCH QUERIES	
Examiner	:	Opsasnick	

APPELLANT'S BRIEF
PURSUANT TO 37 C.F.R. § 1.192

Board of Patent Appeals and Interferences
Washington, D.C. 20231

Dear Sir:

Appellant, Applicant in the above-captioned patent application, appeals the final rejection of Claims 1-55 set forth in the final Office Action mailed on November 26, 2003. A check for the filing fee is enclosed. Please charge any additional fees that may be required now or in the future to Deposit Account No. 11-1410.

I. REAL PARTY IN INTEREST

The real party of interest in the present application is Amazon.com, Inc.

II. RELATED APPEALS AND INTERFERENCES

No related appeals or interferences are pending.

III. STATUS OF CLAIMS

Claims 1-55 are currently pending in the application, and are attached hereto as an appendix. All of the pending claims were finally rejected by the Examiner and are the subject of this appeal.

IV. STATUS OF AMENDMENTS

No claim amendments were made in response to the final Office Action.

V. SUMMARY OF THE INVENTION

The present application discloses, among other things, a query submission/capture process by which a user can efficiently and reliably submit a search query for searching a database of items. The method is particularly well suited for conducting database searches by telephone. In a preferred embodiment, the user is prompted to use a telephone keypad and/or voice to enter a set of characters of the search query (e.g., the first three characters of the query or of a particular term in the query). The user is also prompted to utter the full search query. For example, to search for books by the author "Stephen King," a user may initially press the keys containing the characters "S-T-E" on a telephone keypad (and/or utter these characters), and then utter the name "Stephen King."

To interpret the utterance of the search query (this utterance is also referred to as a "voice query"), the set of characters received from the user is preferably used to dynamically generate a corresponding voice recognition grammar, or to select a previously generated voice recognition grammar from memory. For instance, in the example above, a grammar corresponding to the letters "S-T-E" would be generated or selected from memory, and would then be used to interpret the user's utterance of the name "Stephen King." This voice recognition grammar specifies valid utterances, such as valid words, phrases and names that may be uttered. If the grammar is generated dynamically at the time of the search, it is referred to as a "dynamic grammar."

The voice recognition grammar is preferably derived from a corresponding subset of the items in the domain begin searched (e.g., all items having author names starting with "S-T-E"). As a result, the grammar tends to be much smaller in size than a grammar suitable for searching the entire domain of items. The relatively small size of the grammar improves the reliability of

the voice recognition process, significantly improving the likelihood that the user's utterance of the search query will be accurately captured.

Once a grammar has been generated for a given set of characters, it may be cached in memory and re-used over a period of time.

VI. ISSUES PRESENTED ON APPEAL

As set forth below, Appellant has grouped the pending claims into six claim groups, each consisting of a single independent claim and its respective dependent claims. In view of these groupings, only the following issues are presented on appeal:

Whether independent Claim 1 is properly rejected under 35 U.S.C. § 103 as unpatentable over U.S. Patent 5,917,889 ("Brotman") in view of U.S. Patent No. 6,532,444 ("Weber"), in further view of U.S. Patent No. 5,995,928 ("Nguyen") (collectively "the applied references"); and

Whether independent Claims 15, 24, 33, 43 and 50 are properly rejected under 35 U.S.C. § 103 as unpatentable over Brotman in view of Weber.

VII. GROUPING OF CLAIMS

All of the rejected claims in the present application should not stand or fall together. Appellant, for purposes of this appeal only, is grouping the claims as follows:

GROUP 1: Independent Claim 1 and its dependent claims (collectively Claims 1-14, 39 and 40). These claims are directed generally to an embodiment in which the grammar used to interpret a voice query (i.e., an utterance of a search query) is a "dynamic grammar" that is generated after the user has submitted a set of characters that define a portion of the search query. The dynamic grammar is generated based at least in part on an identified subset of items that correspond to the set of characters received from the user. As set forth in the specification and in some of the dependent claims, the grammar may be generated by extracting selected terms (words, phrases, names, etc.) from the identified subset of items.

GROUP 2: Independent Claim 15 and its dependent claims (collectively Claims 15-23 and 41). These claims are directed generally to a method in which a user's entry of a subset of characters of a query, together with the user's utterance of the full query, are used in combination

to capture the query. The method includes receiving a set of characters entered by a user, where the set of characters represents a portion of the query. In response to receiving the set of characters, a grammar is selected which is derived at least in-part from text extracted from a subset of items that correspond to the set of characters entered by the user. The grammar is provided to a voice recognition system for use in interpreting the query as entered by the user by voice.

GROUP 3: Independent Claim 24 and its dependent claims (collectively Claims 24-32). These claims are directed to a system for conducting searches by voice. The system includes a first code module which causes a user to be prompted to enter a set of characters of a query such that the user may partially specify the query, and a second code module which causes the user to be prompted to utter the query. The system also includes a query server that is programmed to use the set of characters to select a grammar for use by a voice recognition system to interpret the query as uttered by the user.

GROUP 4: Independent Claim 33 and its dependent claims (collectively Claims 33-38 and 42). Claim 33 involves a process by which a user who has submitted a voice query can refine the query by adding an additional query term by voice. The user's utterance of the additional query term is interpreted using a grammar generated at least in part by extracting text from a set of search result items that are responsive to the query.

GROUP 5: Independent Claim 43 and its dependent claims (collectively Claims 43-47). These claims are directed generally to a method for facilitating database searches conducted over a telephone. The method involves interpreting a user's utterance of a search query using a voice recognition grammar that corresponds to a sequence of keys depressed by the user, where this sequence of keys corresponds to a sequence of characters of a query term of a search query.

GROUP 6: Independent Claim 50 and its dependent claims (collectively Claims 50-55). These claims are directed generally to a method which involves receiving from a user an indication, specified at least in part as telephone keypad entries, of a subset of the characters contained in a search query, and also receiving from the user a voice utterance of the entire search query. The voice utterance of the entire search query is interpreted using a voice recognition grammar that corresponds to the indication of the subset of characters.

VIII. ARGUMENT

A. EACH CLAIM GROUP IS PATENTABLY SEPARATE FROM THE OTHERS

Each claim group is patentably separate from the others for the following reasons:

The claims in Group 2 are directed to a method in which a user's entry of a subset of characters of a query, together with the user's utterance of the full query, are used in combination to capture the query. This feature of the Group 2 claims provides a basis for patentability over the applied references.

The claims in Group 1 are directed to an embodiment in which the grammar is a dynamic grammar generated after the user has submitted a set of characters that define a portion of the search query. This feature is not required by the independent claim of Group 2, and provides a separate basis for patentability over the applied references.

Unlike the independent claims in Groups 1 and 2, the claims of Group 3 call for a particular arrangement of components, including a database, a query server, a first code module, and second code module, that collectively operate in a novel way to facilitate database searches by voice. This arrangement of components provides a separate basis for patentability over the applied references relative to Groups 1 and 2.

Unlike the independent claims of Groups 1-3, the claims in Group 4 are directed to a process for refining a search query by adding an additional query term by voice. This process provides a separate basis for patentability over the applied references.

Unlike the independent claims in Groups 1-4, the claims in Groups 5 and 6 involve the combined use of a telephone keypad and voice to specify a query. This combined use of a telephone keypad and voice, as set forth in these claims, provides a separate basis for patentability over the applied references. In addition, Group 5 is patentably separate from Group 6 because Group 5 requires a user to be prompted to depress a sequence of telephone keypad keys corresponding to a sequence of characters of a query term; this aspect of the Group 5 claims is not suggested by the applied references.

B. DISCUSSION OF REFERENCES RELIED UPON BY EXAMINER

In finally rejecting independent Claims 15, 24, 33, 43 and 50, the Examiner relied solely on Weber and Brotman. The Examiner additionally relied on Nguyen in rejecting Claim 1. Each reference is discussed below.

For purposes of this appeal, Appellant will treat Weber and Nguyen as prior art. As these references were not published more than one year before Appellant's filing date, Appellant reserves the right to later disqualify one or both references as prior art.

1. Brotman

Brotman discloses techniques for reliably capturing a string of characters specified by a user over a telephone. These techniques involve having the user both (1) utter all of the characters in the intended string, and (2) select the corresponding keys on the telephone keypad for each of these characters—either by depressing these keys or by uttering the number (0-9) of each such key. The character utterances and the keypad selections are then used in combination to predict the characters intended by the user.

For example, to input the word "cat," the caller would utter the letters C-A-T, and would depress the corresponding telephone keys 2-2-8 (or alternatively, utter the numbers 2-2-8). The keypad selections would then be used to limit the possible interpretations of the character utterances. For example, to interpret the utterance of the letter "C," the system may treat "A," "B" and "C" (corresponding to the key for "2") as the only valid utterances. The likelihood of misrecognition events is therefore reduced. See column 3, lines 42-57 of Brotman. Brotman also discloses the use of grammar prediction rules to predict what the user intended (see column 4, lines 46-59).

Brotman is not directed to the capture of search queries. Even if Brotman's method were used to capture search queries, it would not provide an efficient process for doing so. For example, a user wishing to search for "Stephen King" would apparently have to utter all eleven letters of his name, and would also have to select the corresponding eleven keys on the telephone keypad. In contrast, in Appellant's preferred embodiment, the user could conduct this search by pressing the telephone keys containing the characters "S-T-E" (and/or uttering these characters) and uttering the name "Stephen King." The increase in efficiency is even greater for longer queries.

2. Weber

Weber discloses a user interface through which a user can interact with a computer system by voice. The user's utterances are interpreted in-part using context-specific voice recognition grammars that correspond to specific subjects such as "news," "weather," and "stocks." Col. 6, lines 45-61. Unlike the method disclosed in the present application, these context-specific grammars are not selected based on the user's entry of specific characters of the uttered term or phrase. Rather, they are apparently selected based on the topic or subject currently being browsed, as determined from prior utterances. See col. 3, lines 10-15 and col. 8, lines 1-8.

In contrast to the method disclosed in the present application, Weber's method of using context-specific voice recognition grammars is not well suited for searching large domains of items, such as a domain of millions of book titles or music titles. If Weber's method were used for this purpose, the user would likely have to "drill down" through multiple levels of item categories and subcategories (e.g., books\fiction\mysteries); otherwise, the voice recognition grammars would most likely be too large to provide reliable voice recognition. In addition to being burdensome to users, such an approach would require the users to know how the items they are searching for are categorized. Brotman does not suggest a solution to this deficiency in Weber.

Weber and Brotman do not collectively suggest a method in which a user's entry of a set of characters of a search query is used to select or generate a grammar for interpreting the user's utterance of the query, as set forth in several of the claims at issue. In addition, as discussed below, it would not have been obvious to combine the teachings of Weber and Brotman because, among other reasons, they involve very different types of devices. Specifically, Weber apparently involves interaction with a personal computer, while Brotman involves the entry of character strings by telephone.

3. Nguyen

Nguyen discloses a system that recognizes words as they are continuously spelled by voice by a user. The system includes a speech recognition engine which periodically outputs an updated string of hypothesized letters as the user spells the desired word(s). These letter strings, which may contain errors such as misrecognitions, are passed to a spell checker that repeatedly

compares the strings to a vocabulary list until a best match is found. The matching word may be identified, and presented to the user, before the user has finished spelling the word. See Nguyen at col. 4, lines 26-46.

As with Weber and Brotman, Nguyen does not suggest a method in which a user's entry of a set of characters of a search query is used to select or generate a grammar for interpreting the user's utterance of the query.

C. DISCUSSION OF THE ISSUES ON APPEAL

For the reasons set forth below, Appellant submits that the rejections of the claims in Groups 1-6 are improper.

1. The rejections of the claims in Group 1 are improper because Brotman, Weber and Nguyen do not disclose or suggest all of the limitations of independent Claim 1.

Claim 1 is directed to an embodiment in which the grammar used to interpret a voice query (i.e., an utterance of a search query) is a "dynamic grammar" that is generated after the user has submitted a set of characters that define a portion of the search query. The dynamic grammar is generated based at least in part on an identified subset of items that correspond to the set of characters received from the user.

In rejecting Claim 1, the Examiner appears to take the position that Brotman discloses "identifying a subset of items in the domain [of items being searched] that correspond to the set of characters [received from the user];" and "generating a dynamic grammar based at least in part on the subset of items." Appellants disagree. The portion of Brotman cited in connection with these limitations, namely column 4, lines 46-67, does not disclose or suggest these steps. Although Brotman discloses the generation of a grammar for interpreting a set of character utterances (see col. 4, lines 36-41), this grammar is not generated based at least in part on a subset of items, within a domain of items being searched, that correspond to a set of characters received from the user. Rather, the grammar is generated based directly on the set of telephone keys selected by the user.

Neither Weber nor Nguyen cures this "deficiency" in Brotman. Although Weber discloses the use of a context-sensitive grammar file to interpret a voice query, this context-

sensitive grammar file is a pre-generated grammar file that is selected based on the subject area ("news," "weather," "stocks," etc.) selected by the user. See Weber at col. 6, lines 49-61, and col. 12, lines 12-19. Thus, Weber's context-sensitive grammar file is not a dynamic grammar generated based at least in part on an identified subset of items within the domain being searched.

Brotman, Weber and Nguyen also fail to disclose or suggest a method in which the user submits a "set of characters that define a portion of the query" for purposes of improving recognition accuracy of a voice query. In connection with this aspect of Claim 1, the Examiner cites Nguyen's use of spell checking to predict, and output to the user, a word being spelled by the user. Office Action at page 3, second paragraph. Nguyen, however, does not involve the processing of voice queries, and is not directed to the problem of generating a grammar for interpreting an utterance. Thus, Nguyen cannot reasonably be read as suggesting that a user may be prompted to enter a set of characters that define a portion of search query, or that a submitted set of characters may be used to generate a grammar.

Because Claim 1 includes limitations that are not disclosed or suggested by the applied references, the obviousness rejection of the Group 1 claims is improper and should be withdrawn. As set forth separately below, Appellant additionally submits that the rejection is improper because the Examiner has not identified a suggestion or motivation to combine Brotman and Weber.

2. The rejections of the claims in Groups 2-6 are improper because Brotman and Weber do not disclose or suggest all of the limitations of independent Claims 15, 24, 33, 43 and 50.

The rejections of the claims in Groups 2-6 are improper because, as set forth below, Claims 15, 24, 33, 43 and 50 include limitations that are not disclosed or suggested by Brotman and Weber.

Claim 15

Claim 15 is directed to a method for improving voice recognition accuracy when a user submits a query by voice to search a domain of items. The method comprises "receiving a set of characters entered by a user, the set of characters representing a portion of a query." Neither Brotman nor Weber discloses this step.

Claim 15 also recites "in response to receiving the set of characters, selecting a grammar which is derived at least in-part from text extracted from a subset of items that correspond to the set of characters entered by the user;" and "providing the grammar to a voice recognition system for use in interpreting the query as entered by the user by voice." (Note that the "whereby" clause in Claim 15 clarifies that the phrase "the query as entered by the user by voice" refers to an utterance by the user of the full query.) Neither Brotman nor Weber discloses this combination of limitations. In this regard, as discussed above, Weber selects the context-specific grammars based on the subject or context being browsed, and not based on the user's entry of a set of characters of a query.

Brotman and Weber also fail to disclose or suggest a method in which "the user's entry of a subset of characters of the query, together with the user's utterance of the full query, are used in combination to capture the query."

Because Brotman and Weber fail to suggest all of the limitation of Claim 15, the obviousness rejection of Claim 15 is improper.

In view of the foregoing, Appellant submits that Claim 15, and the claims that depend from Claim 15, are patentably distinct from Brotman and Weber. As discussed separately below, Appellant additionally submits that the rejection is improper because the Examiner has not identified a legally sufficient suggestion or motivation to combine Brotman and Weber.

Claim 24

Claim 24 is directed to a system that includes "a first code module which causes a user to be prompted to enter a set of characters of a query such that the user may partially specify the query," and "a second code module which causes the user to be prompted to utter the query." The claim also calls for a query server that "is programmed to use the set of characters to select a grammar for use by [a] voice recognition system to interpret the query as uttered by the user." As discussed above, Brotman and Weber do not disclose such a system. Claim 24, and the claims that depend from Claim 24, are therefore patentably distinct from Brotman and Weber.

Claim 33

Claim 33 involves a method in which a user refines a search query by uttering an additional query term to add to the query. To interpret the user's utterance of the additional query term, a grammar is generated at least in-part by extracting text from the set of search result

items resulting from the query. A preferred embodiment of this method is described beginning at page 9, line 24 of the present application.

The Examiner has effectively disregarded the query refinement limitations of Claim 33 by grouping it with other independent claims that do not involve search query refinement. Indeed, neither Brotman nor Weber discloses a query refinement process, let alone the particular query refinement process defined in Claim 33. Claim 33, and the claims that depend from Claim 33, are therefore patentably distinct from Brotman and Weber.

Claim 43

Claim 43 is directed to a method that involves "prompting a user to depress a sequence of telephone keypad keys corresponding to a sequence of characters of a query term of a search query." The user is also prompted "to utter the search query by voice." The voice utterance of the search query is interpreted "using a voice recognition grammar that corresponds to the sequence of keys depressed by the user." As discussed above, Brotman and Weber do not disclose or suggest such a method. Claim 43, and the claims that depend from Claim 43, are therefore patentably distinct from Brotman and Weber.

Claim 50

Claim 50 is directed to a method of capturing a search query specified by a user by telephone. The method comprises "receiving from the user an indication of a subset of the characters contained in the search query, said indication of the subset of characters being specified at least in part as telephone keypad entries." The method further comprises "receiving from the user a voice utterance that represents the entire search query," and "interpreting the voice utterance using a voice recognition grammar that corresponds to the indication of the subset of characters."

As discussed above, Brotman and Weber do not disclose or suggest such a method. Indeed, neither reference discloses *any* method of capturing a search query specified by a user by telephone. Claim 50, and the claims that depend from Claim 50, are therefore patentably distinct from Brotman and Weber.

3. **The rejections of the claims of Groups 1-6 are improper because the Examiner has not identified a suggestion to combine Brotman and Weber, and because no such suggestion exists in the references.**

Appellants further respectfully submit that in rejecting the claims of Groups 1-6, the Examiner has failed to identify a legally sufficient suggestion or motivation to combine or modify Brotman and Weber. Appellants further submit that no such suggestion or motivation exists within these references.

The only basis given by the Examiner for combining these two references is that the combination "would make the system taught by Brotman a more interactive speech recognition system." Office Action at page 3, first paragraph and page 6, third paragraph, citing Weber at col. 2, lines 11-50 and 53-65). This statement, however, does not appear to come from the references themselves or from any other identified source of prior art. In addition, it is not clear what the Examiner means by "a more interactive speech recognition system."

Appellant further submits that no suggestion to combine Brotman and Weber exists within the references themselves. In this regard, the two references involve different types of devices: Weber involves user interaction with a personal computer via a microphone, while Brotman involves the entry of character strings by telephone. Given the difference in device types, one skilled in the art would not be motivated from the references themselves to combine their respective teachings.

IX. CONCLUSION

For the reasons set forth above, Appellant submits that the rejections of Groups 1-6 are improper, and requests that the rejection be reversed.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: 5-24-04

By: _____
Ronald J. Schoenbaum, Reg. No. 38,297
Attorney of Record
2040 Main Street, 14th Floor
Irvine, CA 92614

APPENDIX – PENDING CLAIMS

1. A method for improving voice recognition accuracy when a user submits a search query by voice to search a domain of items, the method comprising:

prompting a user to submit a set of characters of a voice query for searching the domain of items, and receiving the set of characters from the user, wherein the voice query is an utterance by the user of a search query, and the set of characters defines a portion of the search query;

in response to receiving the set of characters from the user, identifying a subset of items in the domain that correspond to the set of characters;

generating a dynamic grammar based at least in part on the subset of items, said grammar specifying valid utterances for interpreting the voice query;

prompting the user to submit the voice query, and receiving the voice query from the user; and

interpreting the voice query using the dynamic grammar.

2. The method as in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to submit the first N characters of a query term, where N is greater than 1.

3. The method as defined in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to submit a set of characters of an author's name.

4. The method as defined in Claim 3, wherein generating a dynamic grammar comprises incorporating into the grammar names of authors of the items within the subset of items.

5. The method as defined in Claim 4, wherein the dynamic grammar consists essentially of the names of the authors of the items within the subset of items.

6. The method as defined in Claim 4, further comprising incorporating into the dynamic grammar non-author terms extracted from the subset of items.

7. The method as defined in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to select the characters on a telephone keypad.

8. The method as defined in Claim 7, wherein prompting a user to submit a set of characters further comprises prompting the user to utter the characters, and wherein receiving the set of characters comprises using the keypad entries of the characters to interpret utterances by the user of the characters.

9. The method as defined in Claim 1, wherein generating a dynamic grammar comprises extracting text from the subset of items.

10. The method as defined in Claim 9, wherein extracting text from the subset of items comprises extracting the text from a database field corresponding to a search context of the query.

11. The method as defined in Claim 1, further comprising storing the dynamic grammar within a cache for subsequent use.

12. The method as defined in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to enter a fixed number of characters, wherein the fixed number is selected based on a target grammar size.

13. The method as defined in Claim 1, wherein receiving the set of characters comprises determining in real time whether a number of entered characters is sufficient to produce a grammar that falls below a threshold size.

14. The method as defined in Claim 1, further comprising:

executing a search using the voice query as interpreted using the dynamic grammar to identify a set of search result items;

providing the user an option to add an additional query term to the voice query to refine the search;

generating a second dynamic grammar at least in part from the set of search result items; and

receiving a voice entry of the additional query term from the user, and interpreting the voice entry using the second dynamic grammar.

15. A method for improving voice recognition accuracy when a user submits a query by voice to search a domain of items, the method comprising:

receiving a set of characters entered by a user, the set of characters representing a portion of a query;

in response to receiving the set of characters, selecting a grammar which is derived at least in-part from text extracted from a subset of items that correspond to the set of characters entered by the user; and

providing the grammar to a voice recognition system for use in interpreting the query as entered by the user by voice;

whereby the user's entry of a subset of characters of the query, together with the user's utterance of the full query, are used in combination to capture the query.

16. The method as defined in Claim 15, wherein selecting a grammar comprises:

executing an initial search to identify the subset of items that correspond to the set of characters; and

extracting text from the subset of items for incorporation into the grammar.

17. The method as defined in Claim 16, wherein extracting text from the subset of items comprises extracting the text from a database field corresponding to a search context of the query.

18. The method as defined in Claim 17, wherein the search context comprises an author search, and the database field is an author field.

19. The method as defined in Claim 15, wherein selecting a grammar comprises reading a previously generated grammar from memory based on the set of characters entered by the user.

20. The method as in Claim 15, wherein receiving a set of characters comprises receiving the first N characters of a query term, where N is greater than 1.

21. The method as in Claim 15, wherein receiving a set of characters comprises receiving characters entered at least in-part using a telephone keypad.

22. The method as in Claim 15, wherein receiving a set of characters comprises using a telephone keypad entry of a character by the user to interpret an utterance of the character by the user.

23. (original) The method as defined in Claim 15, wherein receiving a set of characters comprises determining in real time whether a number of entered characters is sufficient to produce a grammar that falls below a threshold size.

24. A system for conducting searches by voice, comprising:

a database of items;

a query server which searches the database of items according to voice queries from users, the query server coupled to a voice recognition system which interprets the voice queries according to grammars;

a first code module which causes a user to be prompted to enter a set of characters of a query such that the user may partially specify the query; and

a second code module which causes the user to be prompted to utter the query;

wherein the query server is programmed to use the set of characters to select a grammar for use by the voice recognition system to interpret the query as uttered by the user.

25. The system as defined in Claim 24, wherein the first and second code modules comprise voiceXML coding.

26. The system as defined in Claim 24, wherein the query server selects the grammar by at least:

executing a preliminary search to identify a subset of items that match the set of characters; and

extracting text from the subset of items to incorporate into grammar.

27. The system as defined in Claim 26, wherein the query server is programmed to extract author names from the subset of items to generate a grammar for performing a voice-based author search.

28. The system as defined in Claim 24, wherein the query server is programmed to select the grammar from memory using the set of characters.

29. The system as defined in Claim 24, wherein the set of characters is a set of the first N letters of a query term, where N is greater than 1.

30. The system as defined in Claim 29, wherein the query term is a name of an author.

31. (original) The system as defined in Claim 29, wherein N is selected based on a target grammar size.

32. The system as defined in Claim 24, wherein the first code module prompts the user to both utter, and enter on a telephone keypad, each alphabetic character of the set.

33. A method of assisting users in locating items in a database using voice queries, the method comprising:

receiving a voice query from a user, and identifying a set of search result items that are responsive to the voice query;

providing the user an option to refine the query by adding an additional query term;

generating a grammar at least in-part by extracting text from the set of search result items; and

using the grammar to interpret an utterance by the user of an additional query term.

34. The method as defined in Claim 33, wherein generating a grammar comprises extracting text from a database field corresponding to a search context of the query.

35. The method as defined in Claim 33, wherein using the grammar to interpret an utterance comprises using the grammar to interpret utterances of multiple additional query terms by the user.

36. The method as defined in Claim 33, wherein the grammar is generated in response to selection by the user of the option to add an additional query term.

37. The method as defined in Claim 33, wherein the option to refine the query is presented to the user only if the number of items in the set exceeds a predefined threshold.

38. The method as defined in Claim 33, further comprising storing the grammar in a cache for use with subsequent query submissions.

39. The method as in Claim 1, wherein the set of characters is a subset of the characters contained in a textual representation of the voice query.

40. A system that operates according to the method of Claim 1.

41. A system that operates according to the method of Claim 15.

42. A system that operates according to the method of Claim 33.

43. A method for facilitating database searches conducted over a telephone, the method comprising:

prompting a user to depress a sequence of telephone keypad keys corresponding to a sequence of characters of a query term of a search query, and identifying a resulting sequence of keys depressed by the user;

prompting the user to utter the search query by voice, and receiving a resulting voice utterance from the user; and

interpreting the voice utterance using a voice recognition grammar that corresponds to the sequence of keys depressed by the user, said voice recognition grammar specifying valid utterances.

44. The method of Claim 43, wherein the search query consists of said query term.

45. The method of Claim 43, wherein the search query contains multiple query terms.

46. The method of Claim 43, further comprising prompting the user to utter said sequence of characters by voice, and using resulting voice utterances of the characters in combination with the sequence of keys depressed by the user to identify the sequence of characters intended by the user.

47. The method of Claim 43, further comprising selecting the voice recognition grammar from a repository of previously-generated voice recognition grammars in which different voice recognition grammars correspond to different sequences of characters.

48. The method of Claim 43, further comprising generating the voice recognition grammar on-the-fly based on input from the user.

49. A system that operates according to the method of Claim 43.

50. A method of capturing a search query specified by a user by telephone, the method comprising:

receiving from the user an indication of a subset of the characters contained in the search query, said indication of the subset of characters being specified at least in part as telephone keypad entries;

receiving from the user a voice utterance that represents the entire search query; and

interpreting the voice utterance using a voice recognition grammar that corresponds to the indication of the subset of characters, said voice recognition grammar specifying valid utterances.

51. The method of Claim 50, wherein the indication of the subset of characters further comprises respective voice utterances of the characters in the subset.

52. The method of Claim 50, further comprising selecting the voice recognition grammar from a repository of previously-generated voice recognition grammars in which different voice recognition grammars correspond to different sets of characters.

53. The method of Claim 50, further comprising generating the voice recognition grammar on-the-fly in response to input from the user.

54. The method of Claim 50, further comprising executing a database search using a textual representation of the voice utterance.

55. A system that operates according to the method of Claim 50.